

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-044126

(43)Date of publication of application : 08.02.2002

(51)Int.Cl.

H04L 12/56

(21)Application number : 2000-220617

(71)Applicant : NEC CORP

(22)Date of filing : 21.07.2000

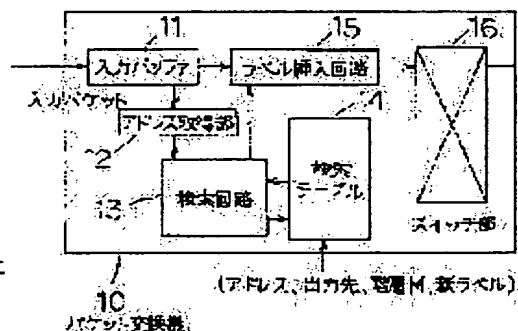
(72)Inventor : UENO YOJI

(54) PACKET TRANSMISSION METHOD AND UNIT

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a packet transfer unit that reduces a processing load at node-ware transfer.

SOLUTION: An input buffer 11 temporarily stores a received packet and an address acquisition section 12 acquires information required for transfer. A retrieval circuit 13 uses the extracted destination address to retrieve a retrieval table 14. The retrieval circuit 13 uses the destination address for a key to extract output information. In this case, a maximum stack layer number M with respect to a packet group having a certain destination is designated. A packet capsulated by a label insertion circuit 15 is transferred to a desired output destination port via a switch section 16.



LEGAL STATUS

[Date of request for examination] 19.06.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3543952

[Date of registration] 16.04.2004

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2002-44126
(P2002-44126A)

(43) 公開日 平成14年2月8日 (2002.2.8)

(51) Int.Cl.⁷
H 0 4 L 12/56

識別記号

F I
H 0 4 L 11/20

タームコード* (参考)

1 0 2 A 5 K 0 3 0

審査請求 有 請求項の数 5 O L (全 6 頁)

(21) 出願番号 特願2000-220617 (P2000-220617)

(22) 出願日 平成12年7月21日 (2000.7.21)

(71) 出願人 000004237

日本電気株式会社
東京都港区芝五丁目7番1号

(72) 発明者 上野 洋史

東京都港区芝五丁目7番1号 日本電気株式会社内

(74) 代理人 100088328

弁理士 金田 暢之 (外2名)

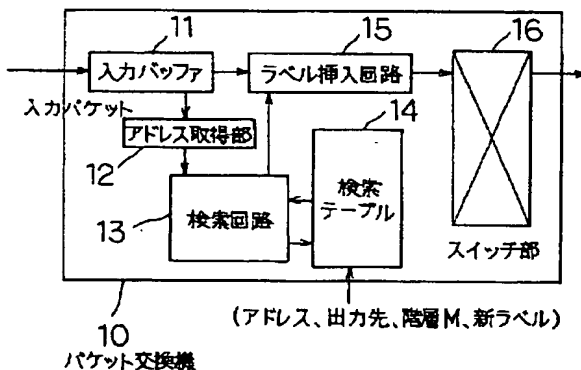
Fターム(参考) 5K030 GA03 HA08 JA05 KX12 KX24
KX27 LB05 MA14

(54) 【発明の名称】 パケット転送方法および装置

(57) 【要約】

【課題】 ノードウェア転送における処理負荷を軽減する。

【解決手段】 入力パケットは入力バッファ11に一時格納され、アドレス取得部12で転送に必要な情報を取り出す。ここで取り出した宛先アドレスを使用して検索回路13にて検索テーブル14を検索する。検索回路13では宛先アドレスをキーとして出力に関する情報を取り出す。ここで、ある宛先を持つパケット群に対する最大のスタック階層数Mを指定しておく。ラベル挿入回路15にてカプセル化されたパケットはスイッチ部16を経て所望の出力先ポートへと転送される。



【特許請求の範囲】

【請求項1】 パケット転送方法において、設定に基づいて宛先毎に決められた最大ラベル階層数のラベルスタックエントリーを、パケット転送網への入り口エッジノードで付加することを特徴とするパケット転送方法。

【請求項2】 複数階層のラベルスタックエントリーの先頭ラベル位置を、ポインタビットを使用して指示し、また、このポインタビット値を増減させることでスタックの追加／削除と等価な処理を行う、請求項1記載のパケット転送方法。

【請求項3】 入力パケットを一時格納する入力バッファと、
前記入力パケットのヘッダ部から宛先アドレスを取り出すアドレス取得部と、
宛先アドレスと出力先ポート、付加するラベルスタックエントリーの階層数、付加する新ラベルを含む出力情報の対応が記述されている検索テーブルと、
前記アドレス取得部で取得された宛先アドレスを用いて前記検索テーブルを検索し、前記出力情報を得る検索回路と、
前記検索で得られた出力情報で指示されるラベルスタックエントリーをパケットの先頭に付加するラベル挿入回路と、
所望の出力先ポートへ出力情報を振り分けるためのスイッチ部を有するパケット交換機。

【請求項4】 入力パケットを一時格納する入力バッファと、
入力パケットのラベルスタックエントリーを取り出すラベルスタック取得部と、
取り出したラベルスタックエントリーを使用して処理判定を行う第1のスタック処理回路と、
出力先ポート、ラベルに対する処理、新ラベルが記述されている検索テーブルと、
入力ラベルを用いて前記検索テーブルを検索する検索回路と、
前記検索結果を用いてスタック操作を行う第2のスタック処理回路と、
所望の出力先ポートへ出力情報を振り分けるためのスイッチ部を有するパケット交換機。

【請求項5】 前記第1のスタック処理回路は、入力されたパケットの先頭ラベルスタックエントリー中のポインタビット値を参照し、Sビットよりラベルの階層数を取得し、前記ポインタビット値が指し示す位置のラベルを取り出し、該ラベルをキーとして前記検索回路により前記検索テーブルを検索し、処理内容を判定し、
前記第2のスタック処理回路は、処理内容がN階層のラベルスタックエントリー追加、すなわちN階層のプッシュ処理の場合、前記ポインタ値がNよりも大きければ新たなラベルスタックエントリーを追加し、ポインタ値をNだけ減少させ、新規スタック位置を参照し、該新規ス

タック位置に新ラベルを付加し、出力パケットを形成し、処理内容がN階層のラベルスタックエントリー削除、すなわちN階層のポップ処理の場合、前記ポインタ値をNだけ増加させ、参照すべきスタック位置を変更し、前記スタック値が全階層数Mより大きいとき、ベイロードに付加されているラベルスタックエントリーを全て削除し、出力パケットを形成し、処理内容がラベル交換の場合には、前記ポインタビット値により指示される位置のラベルを新規ラベルに交換し、出力パケットを形成する、請求項4記載のパケット交換機。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はパケット転送方法および装置に関する。

【0002】

【従来の技術】近年注目されているIPトラヒックの転送方式としてMPLS（マルチプロトコルラベルスイッチング）転送方式がIETF（Internet Engineering Task Force）にて標準化検討されている。MPLSではラベルスイッチングルータ間の転送方路決定をラベルと呼ぶ固定長の数値を参照することにより行う。

【0003】MPLSの特徴の一つとしてラベルの階層化が挙げられる。すなわち、ラベルを階層化し、それぞれを個別に管理することにより、あるラベル付きパケットを網内でトンネリングすることができる。つまり、これまで付いていたラベル値を保持したまま、新規ラベルでカプセル化し、この先頭のラベル値を参照することで網内を転送する。このようにすることにより、網内での経路設定を階層毎に独立に設定することが可能になる。

【0004】このようにラベルを階層化することによる利点は大きいものの、ラベルを付加するためのカプセル化やカプセル分離する際にはパケットのヘッダに対してラベルの挿入、削除という処理が必要である。この処理のためには、待ち合わせのためのバッファが必要であり、また、挿入、削除、等の処理が入力ラベル毎に異なり、このためハードウェアでスイッチングするルータでの処理には不向きである面を持ち合わせている。

【0005】MPLS標準では入力されたラベルの先頭ラベルを参照して出力処理を決定する。出力処理にはラベル付け替え、ラベルスタックの削除（ポップ）、ラベルスタックの追加（プッシュ）がある。ラベル付け替えは入力ラベルから検索した結果の出力ラベルを入力ラベルに置き換える処理である。ラベルスタックの削除は、先頭ラベルを削除する処理であるが、その結果、得られる先頭ラベル、あるいはカプセル化されていたIPパケットのIPアドレスを使用して再度検索する必要がある。つまり、これらの検索をして最終的な出力が決定されるまで、パケットはバッファに格納しておかなければならない。また、ラベルスタックの追加は、これまでのラベル位置の先頭にさらにラベルエントリーを付加する処理で

ある。この場合も新規ラベルを挿入する時間だけ、パケットの送信を待ち合わせる必要があり、バッファにおいて待ち合わせ処理が発生する。

【0006】このように、それぞれの処理は簡単であるが、ハードウェアルータ（パケット交換機）で処理を実現する場合には処理が複雑になり、また、特にラベルスタックのポップ動作において処理時間がかかる、という問題があった。

【0007】本発明の目的は、ハードウェア転送時における処理負荷を軽減することができるパケット転送方法および装置を提供することにある。

【0008】

【課題を解決するための手段】本発明のパケット転送方法は、設定に基づいて宛先毎に決められた最大ラベル階層数のラベルスタックエントリーを、パケット転送網への入り口エッジノードで付加することを特徴とする。また、複数階層のラベルスタックエントリーの先頭ラベル位置を、ポインタビットを使用して指示し、また、このポインタビット値を増減させることでラベルスタックエントリーの追加／削除と等価な処理を行う。

【0009】複数ラベルをパケット交換網の入り口エッジで付加し、網内での転送はポインタビットの値が指示するラベルを参照して決定する。ラベルスタックエントリーの削除／追加処理の場合にはポインタビット値を増減させることにより実現する。このように転送ノードのラベルスタックエントリーの削除／追加処理の際にはポインタビット値のみを変化させ、物理的なラベルスタックエントリーの削除や追加挿入を行わない。

【0010】本発明では予めネットワーク管理端末からの設定、あるいは交換機間のプロトコルのやり取りによって既知である宛先に対する最大階層数を網への入力ノードに設定し、カプセル化する際に最大数のラベルスタックエントリーを挿入することを特徴とする。

【0011】また、この際にどのスタック位置が先頭であるかを示すために、ポインタビットを使用して、参照すべきラベル位置を指定することを特徴としている。中間ノードではこのポインタビットの値を増減することによりラベルの挿入／削除と等価な処理を行う。

【0012】

【発明の実施の形態】次に、本発明の実施の形態について図面を参照して説明する。

【0013】図1に本発明によるラベル付きパケットの構成を示す。ラベル付きパケットはペイロード（被カプセル化パケット）部分200と、その先頭に付加される複数階層のラベルスタックエントリーとから構成される。図1はラベルが4階層（201～204）である場合を示している。パケットの先頭位置のスタックエントリー中に書かれたP（ポインタ）値が3であるとき、有効ラベルの先頭がパケットの先頭から3段目であること、すなわちラベルエントリー202（ラベル#2）で

あることを示す。また、最後尾ラベルスタックエントリーのSビットは1であり、このSビットによりラベルスタックエントリーの最後尾であること（すなわち図1ではエントリー数が4段であることを）を示している。

【0014】図2にラベルスタックエントリーの定義を示す。図2（1）はIETF（Internet Engineering Task Force）において検討されているMPLS（Multi Protocol Label Switching）で定義されているラベルスタックエントリーの図である。32ビットのエントリーに対して、ラベルが20ビット、実験用の予備領域（EXP）が3ビット、スタックの最後尾を示すSビットが1ビット、TTL（Time to Live）値を保持するフィールドが8ビットである。これに対して図2（2）が本発明によるラベルスタックエントリーの定義である。図2（1）との相違はEXP領域をポインタ値Pとして定義し使用することである。Pは先頭から何階層目がラベルスタックの先頭であるかを示すために使用する。

【0015】図3に本発明によるパケット交換機の入力エッジノードの構成を示す。パケット交換機10は入力パケットを一時格納する入力バッファ11と、入力パケットのヘッダ部から宛先アドレスを取り出すアドレス取得部12と、宛先アドレスと出力先ポート、付加するラベルスタックエントリーの階層数、付加する新ラベルを含む出力情報の対応が記述されている検索テーブル14と、その宛先アドレスを用いて検索テーブル14を検索し、出力情報を得る検索回路13と、検索で得られた出力情報で指示されるラベルスタックエントリーをパケットの先頭に付加するラベル挿入回路15と、所望の出力先ポートへ出力情報を振り分けるためのスイッチ部16から構成される。検索テーブル14には入力パケットの宛先アドレスから出力情報を得るためのエントリーが格納されている。出力情報としては出力先ポート、付加するスタックエントリーの階層数M、付加する新ラベル（複数のエントリー可能）が記述されている。例えば2段のラベルスタックエントリーを書き加える場合には2種類の新ラベルを使用する。

【0016】図4に本発明によるパケット交換機の間接ノードの構成を示す。パケット交換機20は入力パケットを一時格納する入力バッファ21と、入力パケットのラベルスタックエントリーを取り出すラベルスタック取得部22と、取得したラベルスタックエントリーを使用して処理判定を行うスタック処理回路23Aと、入力ラベルを用いて検索テーブル25を検索する検索回路24と、検索結果に基づいてラベル交換等のスタック操作を行うスタック処理回路23Bと、所望の出力先ポートへ出力を振り分けるためのスイッチ部26から構成される。検索テーブル25には入力パケットのラベル値から出力情報を得るためのエントリーが格納されている。出力情報としては出力先ポート、ラベルに対する処理（ラベル交換／ラベルプッシュ／ラベルポップ）、新ラベル

等が記述されている。

【0017】図6に本発明によるパケット交換網でのパケット転送の概念図を示す。図6に示すパケット交換網はIPパケット等の被転送パケットの入力ノードであるパケット交換機1、中間ノードであるパケット交換機2～5、被転送パケットの出力ノードであるパケット交換機6から構成される。交換機1への入力パケット100に対して、それぞれの交換機1～6の出力パケットを101～106として図示している。

【0018】図3に示すパケット交換機（入力エッジノード）10の処理を説明する。入力パケットは入力バッファ11に一時格納され、アドレス取得部12で転送に必要な情報を取り出す。入力パケットがIPパケットの場合は、IPヘッダの宛先IPアドレスを取り出す。このとき転送に使用する情報としてさらに細かな情報（送信元IPアドレス、TCP/UDPのポート番号等）を使用することもできる。ここで取り出した宛先アドレスを使用して検索回路13にて検索テーブル14を検索する。検索回路13では宛先アドレスをキーとして出力に関する情報を取り出す。出力情報は予め装置の管理情報、あるいはラベル割り当てのプロトコル等により決定され、検索テーブル25に格納されている。ここで重要なのは、ある宛先を持つパケット群に対する最大のスタック階層数Mを指定しておくことである。この最大スタック数Mに基づいて、入力パケットの先頭部にM階層分のラベルスタックエントリーを図1のように挿入する。図1ではM=4の場合を示している。出力パケットに付加するラベルが1種類の場合は、階層M（最も最後尾のラベル）のエントリーに新規ラベル値を書き込み、階層1（先頭のラベル）のエントリーのP値をMにする。出力パケットに付加するラベルが2種類の場合は、階層Mと階層M-1のエントリーに新規ラベル値をそれぞれ書き込み、階層1のエントリーのP値をM-1にする。ラベル挿入回路15にてカプセル化されたパケットはスイッチ部16を経て所望の出力先ポートへと転送される。

【0019】図4に示すパケット交換機（中間ノード）20の処理を説明する。入力パケットは入力バッファ21に一時格納され、ラベルスタック取得部22で転送に必要な情報を取り出す。ここではパケットの先頭に付加されているラベルスタックエントリーを取り出す。取り出したラベルスタックエントリーに対する処理はスタック処理回路23Aおよび23Bで行われる。ここでの処理については後述する。スタック処理回路23A、23Bで参照すべきであると判定されたラベルをキーとして検索回路24、検索テーブル25において検索処理を行い、出力に関する情報を取り出す。出力に関する情報はラベルに対する処理（付け替え／プッシュ／ポップ）、新規ラベル、出力ポート等がある。スタック処理回路23A、23Bにてラベル操作されたパケットはスイッチ部26を経て所望の出力ポートへと転送される。

【0020】スタック処理回路23A、23Bの内部処理（パケット処理）について、図5に示すフローチャートを用いて説明する。パケット交換機に入力されたパケットは、はじめに先頭ラベルスタックエントリー中のP値を参照する（ステップ201）。また、Sビットが1となっているラベルスタックを判定し、ラベルの階層数Mを取得する（ステップ202）。次いで、Pの値が指し示す位置のラベルを取り出し、このラベルを使用して検索を行う（ステップ203）。ラベル検索の結果、ラベルスタックエントリーの追加（N階層のプッシュ）処理、ラベルスタックエントリー削除（N階層のポップ）処理、ラベル交換処理のいずれかの結果が得られる。

【0021】ラベルスタックエントリー追加（階層の尾プッシュ）処理の場合をステップ204～208に示す。この場合、ラベル検索の結果として、N階層のプッシュ動作であること、出力時の新ラベル、出力先のインタフェース等の情報が得られる。まず、P値がNよりも大きいことを確認する。本発明のパケット転送方法において最大階層数を宛先毎に予約しておくので、この条件になることは通常あり得ないが、もしもP値がNよりも小さい場合は、新たにラベルスタックエントリーを追加する（ステップ208）。ステップ205のチェックの後、P値をNだけ減少させ、新規スタック位置を参照する（ステップ206）。ここで得られた新規のP値が指し示す位置のスタックに新ラベルを付加し（ステップ207）、出力パケットを形成する。

【0022】スタック削除（N階層のポップ）処理の場合をステップ201～212に示す。この場合、ラベル検索の結果として、N階層のポップ動作であること、出力先のインタフェース等の情報が得られる。ラベルスタックエントリーを削除するために、P値をNだけ増加させ、参照すべきスタック位置を変更する（ステップ210）。これによりP値が新スタック位置を指し示すようになるが、Pが全階層数Mよりも大きくなるとき（ステップ211）には、ラベルスタックエントリーが不要になるので、ペイロードに付加しているラベルスタックエントリーを全て削除する（ステップ212）。

【0023】ラベル交換の場合にはP値により指し示される位置のラベルを新規ラベルに交換し、出力パケットを形成する（ステップ213、214）。

【0024】図6に示すパケット交換網の概念図を用いて、本発明によるラベル付け替えを網上で転送例を挙げて説明する。パケット交換機1において、入力パケット（例えばIPパケット）はその宛先アドレス検索結果によって指定されるカプセル化処理を行う。図6においては入力パケット100に対して2階層のラベルスタックエントリーを付加し、最後尾位置である階層1に新ラベルを割り当て、P=2として階層1が先頭ラベルであることを示す。パケット交換機2においてはP値が示すラベル値（P=2；階層1）を使用して検索し、宛先がパケ

ット交換機3であること、ラベルの階層を1段追加(1段プッシュ)することを得る。ここでP値を1だけ減算し、階層2(先頭位置)に新ラベルを割り当てるとともにP=1とする。パケット交換機3、4では階層2のラベルを参照してラベル検索を行い、それぞれ新ラベル値への付け替え処理を行う。パケット交換機5では同様に階層2のラベル検索を行い、その結果、ラベルの階層を1段削除(1段ポップ)することを得る。そこで、P値を1だけ増加させ、P=2とする。これにより、次段の交換機6において階層1のラベルを参照する。この交換機における処理がラベル階層を1段削除(1段ポップ)である場合には、ラベルによるカプセル化削除を意味するので、2階層のラベルスタック領域を削除し、パケット転送網から出力する(パケット106)。

【0025】図7を用いて、パケット交換機1(入力エッジノード)において、入力IPパケットに対する出力先によって付加するスタック階層数が異なる場合を説明する。パケット交換機の回線として入力回線1-1と、出力回線1-2、1-3、1-4を考える。入力回線1-1からの入力パケット100をカプセル化する際に、宛先IPネットワークアドレスが1.2.0.0で、ネットマスクが255.0.0の場合、すなわち1.2.0.0/16で示されるIPネットワークアドレスであるなら、出力回線を1-2とし、付加するスタック数M=2とする。宛先が1.3.0.0/16で示されるIPネットワークアドレスであるなら、出力回線1-3で付加するスタック数M=4、宛先が1.4.0.0/16で示されるIPネットワークアドレスであるなら、出力回線は1-4で付加するスタック数M=3である、というように変化させる。このように、予め管理系端末からの設定や、あるいは交換機間の制御プロトコル等に基づいて、宛先に応じた最大階層数Mを可変とすることができる。これにより、回線の帯域の浪費を抑えることが可能になる。

【0026】本発明によるパケット転送方法では、各パケット交換機においてポインタビットの解釈が必要である。そのため従来のMPLS転送ノードとの境界ノード(パケット交換機)においては、未使用のスタックを取り扱う処理を行う。すなわちポインタビットPがNのとき、先頭からN-1までのラベルスタックエントリは未使用なのでこれらのエントリフィールドを削除することにより、MPLS標準の転送ノードを使用して転送することが可能である。

【0027】また、MPLS転送ノードから、本方式による

パケット転送網へ入力する場合には、先頭のラベルスタックエントリのEXPビットをPと読み替え、P=1を代入する。このとき、予めM値がわかっているならば、空きのラベル階層化エントリを付加する、すなわち網への入力エッジノード(図3におけるパケット交換機10)と同様に処理することが可能である。

【0028】

【発明の効果】以上説明したように本発明によれば、宛先に応じて予め指定される階層数のラベルをエッジノードに付加し、パケット交換網内での転送時にはエントリに割り当てら領域は常に一定で、ラベルエントリ領域の追加、削除を行わないため、ハードウェア転送時において処理負荷を軽減することができる。

【図面の簡単な説明】

【図1】本発明によるラベル付きパケットの構成を示す図である。

【図2】ラベルスタックエントリの構成図である。

【図3】パケット交換機(入力エッジノード)の構成図である。

【図4】パケット交換機(中間ノード)の構成図である。

【図5】スタック処理回路の処理のフローチャートである。

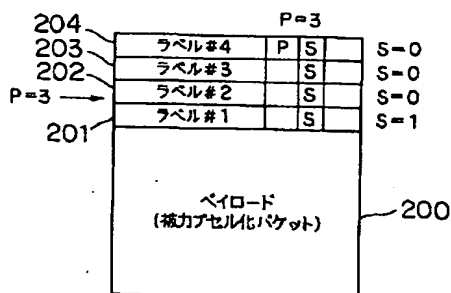
【図6】パケット転送の概念図である。

【図7】階層数Mを変更する例を示す図である。

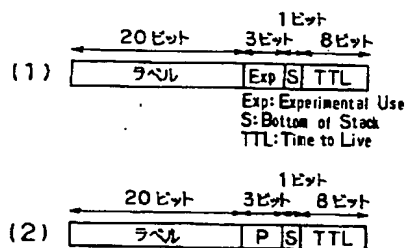
【符号の説明】

200 ベイロード部分
201~204 ラベルスタックエントリ
11、21 入力バッファ
12 アドレス取得部
13、24 検索回路
14、25 検索テーブル
15 ラベル挿入回路
16、26 スイッチ部
22 ラベルスタック取得部
23A、23B スタック処理回路
201~214 ステップ
1~6 パケット交換機
100~106 パケット
1-1 入力回線
1-2、1-3、1-4 出力回線
110、120、130 カプセル化パケット

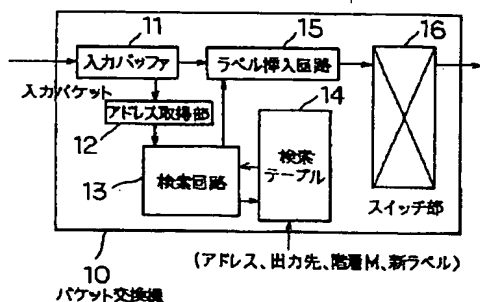
【図1】



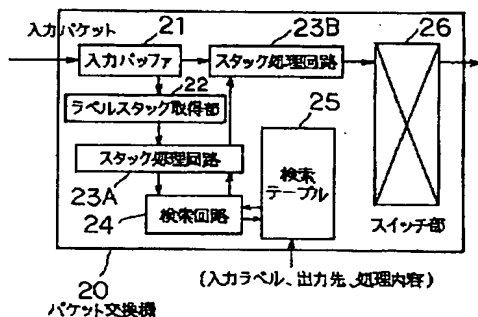
【図2】



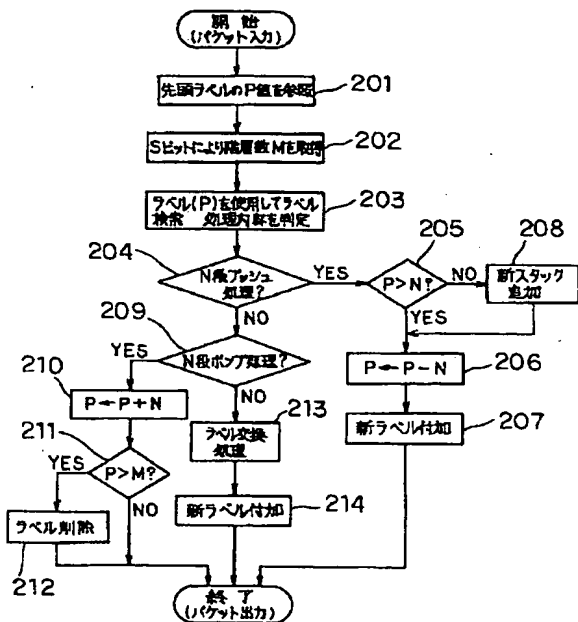
【図3】



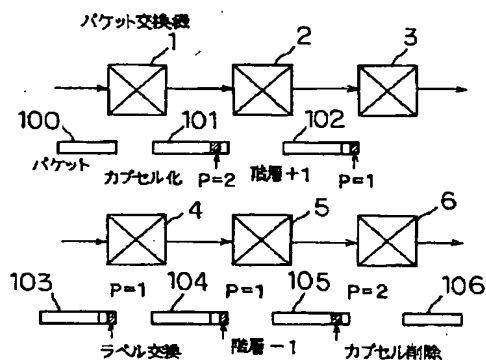
【図4】



【図5】



【図6】



【図7】

